



Questions & Answers Session 2

Please type your questions in the Question Box. We will try our best to get to all your questions. If we don't, feel free to email Erika Podest (erika.podest@jpl.nasa.gov), Amber McCullum (amberjean.mccullum@nasa.gov) or Juan Torres-Peréz (juan.l.torresperez@nasa.gov).

Question 1: Is SAR Imagery good for monitoring forest sanitation?

Answer 1: I'm not sure what is meant by forest sanitation.

Question 2: I can't wait for the Google earth engine, however my google earth engine login details won't work. can you guys help with GEE login??

Answer 2: We cannot help with getting you set up with a GEE account. You can check the GEE help webpage to see how you can get assistance for your login:

<https://developers.google.com/earth-engine/help>

Question 3: Will Sentinel Toolbox load on 32-bit Windows machines?

Answer 3: It does. The following page has the system requirements for the Sentinel Toolbox:

<http://step.esa.int/main/download/snap-download/>

Question 4: As you mentioned supervised and unsupervised classification algorithm, how's about hybrid approach in comparison with the two mentioned methods? Could it yield a better result?

Answer 4: You can run a hybrid approach, which is especially helpful when you do not have good reference data or are not sure what the different class clusters in your image might be. The unsupervised classification will identify clusters of similar backscatter, which you can then name based on your knowledge of backscatter characteristics (e.g. very bright areas are most likely inundated vegetation or urban areas) together with Google Earth images that can help identify what is there. You can then use this information to train a supervised classification. I am not sure whether an approach like this will yield better results especially if you already have good reference data. Fortunately, with GEE it is easy to run different classifications and you can easily test this hybrid approach and compare it to a supervised classification approach only.



We have highlighted in this demonstration an example of using multiple types of data in hope to improve the classification accuracy. There are also techniques that address spatial dependence, such as the impact of topography on land cover types. This could include adding something like a DEM to your classification algorithm. There are also hybrid unsupervised/supervised approaches that use machine learning. Here is a nice review article of some of these approaches, most of which focus on optical imagery classification techniques:

<https://www.tandfonline.com/doi/pdf/10.5721/EuJRS20144723>

Question 5: Which one (method???) is recommended in the area of atmospheric elements classification into homogeneous regions such as rainfall regions?

Answer 5: I'm not sure I understand. For regional classifications, you could look at greener regions/high NDVI regions, which could be an indicator of a region with higher rainfall.

Question 6: Can forest monitoring be done with dualpol? I am finding that quadpol is great for forest cover classifications but it is more temporally limited than dualpol.

Answer 6: Yes, forest monitoring can be done with dual pol and even single pol. Most freely available SAR data is dual pol. In this demo we used VV and VH.

Question 7: Do we need to worry about topographic corrections for radar data?

Answer 7: Yes. In the case of the Sentinel-1 radar data on GEE, it has already been terrain corrected. If you are processing Sentinel-1 data downloaded from say ASF, then you will have to apply a terrain correction using the Sentinel-1 toolbox. ASF also has L-band data from Palsar and most of that data has already been terrain corrected, which is the RTC (radiometrically terrain corrected) data.

Question 8: Can we import training sites to GEE for land classification collected by GPS from the field? If so, how do we import the training sites?

Answer 8: Yes you can. You have to go to "Assets" in the upper left window and select the red "New" button. A menu will open indicating the type of files that you can upload, which includes vector files and tables (.csv). Session three of this webinar will show how to upload a raster and vector file.



Question 9: One of the limitations of optical is the inability of penetrating smoke and cloud. But can't we eliminate this limitation by taking Landsat images taken in the dry season?

Answer 9: Good question, yes clouds and smoke influence the optical imagery, so it is recommended to use images with low to no cloud cover. Using imagery from the dry season, or anytime when cloud cover is a great idea. In some regions it might not be possible to obtain a cloud-free image over the course of the year, so that is where radar can be helpful. If you are interested in monitoring something like agriculture, you may need to take into account the seasonality (green-up and senescence), which means that you'd need a denser temporal sample and sometimes this might be difficult to obtain with optical imagery because of cloud cover.

Question 10: It has been stated that Radar can penetrate through the soil. So, can Radar help us to know the different types of soil? If so, how?

Answer 10: This is very difficult because the radar signal response to different types of soils is very small as compared to the signal response to soil moisture or vegetation above the soil. In order to do this ideally you'd have quad-pol and little to no vegetation. Here are some relevant publications:

<https://www.mdpi.com/2072-4292/7/8/10098/htm>

https://www.researchgate.net/publication/275730910_Digital_Mapping_of_Soil_Texture_Using_RADARSAT-2_Polarimetric_Synthetic_Aperture_Radar_Data

https://www.researchgate.net/publication/254063230_Soil_Texture_Estimation_Over_a_Semiarid_Area_Using_TerraSAR-X_Radar_Data

<https://www.tandfonline.com/doi/full/10.1080/10106049.2019.1618924?scroll=top&neededAccess=true>

Question 11: In slide 21: Could you explain a little more when Radar which is sensitive to structure and the topographic differences can cause distortions in data? I have that doubt because a structure can be natural, comes from the topography. So both ideas can exchange and cause confusion.

Answer 11: Because SAR is a side looking sensor, areas of topography can cause distortions known as layover or foreshortening. The first session in this series discussed these types of distortions. Complex topography can also cause radiometric



distortions where slopes facing the radar appear very bright and can create confusion when identifying classes that have high backscatter such as inundated vegetation or urban areas.

Question 12: Why not Sentinel-2 and Sentinel-1. What motivates the use of Landsat 8? Knowing the difference in both temporal and spatial resolution?

Answer 12: Sentinel-2 is a great option for optical data and could be used in a similar manner to Landsat 8 in this example. We simply choose Landsat 8 for this example. Things to consider when choosing imagery are: what is the time period of available data, what is the spatial and spectral resolution, and what is the repeat time of the sensor. If you are looking back in time, Landsat can also be useful since we have such a long period of record.

Question 13: Why is descending pass important? What is the purpose of filtering by descending orbital pass?

Answer 13: In this case there were more images in the descending pass. But if you are not finding what you need in the descending pass, look for ascending. It is important to remember that you should NEVER combine descending and ascending in a classification or when applying a threshold.

Question 14: Radar is highly affected by topography. Hence, is Radar data applicable to make land cover classification in areas with undulating topographies?

Answer 14: As long as you train your algorithm properly you can apply a landcover classification in areas where there is undulating topography.

Question 15: If you are in an area where terrain creates a shadow, how would you automatically delineate shadow in SAR. You can identify in Landsat 8, but wondering if you could do that in SAR.

Answer 15: You can do that using a DEM. You identify areas with very low backscatter that are within a given slope/elevation range of your DEM and create a mask.

Question 16: Are there any radar datasets or systems available for the 1980s or 1990s to create a time series from 1980s until now?

Answer 16: Refer to the Alaska Satellite Facility website. JERS-1 is an L-band SAR sensor from the Japanese Space Agency that flew in the 90s and there is data over



selected areas. ERS is a C-band sensor from the European Space Agency that flew in the 90's and 2000's and there is data over select areas.

Question 17: While setting the dates for filtering, do we need to first check the availability of data for the dates we select or is data available at any date in all areas?

Answer 17: Remember S-1 has a 12 day repeat pass. Setting the filter dates is a trial and error process. You can start by setting a range of dates to see what is available and then adjust that range accordingly.

Question 18: Can we be made a supervised classification with an averaged image from a temporal interval?

Answer 18: If you are just interested in mapping landcover then creating an averaged image from a temporal interval is fine - assuming that the temporal interval is short or that things have not changed significantly within that time period. The advantage to doing this is that you reduce speckle without reducing the spatial resolution of the image. However, if you are interested in identifying landcover change then doing an average from a time series will cause a loss in information in landcover change and hence this approach is not recommended when looking at change.

Question 19: How can we add other DEM's of better resolution than SRTM?

Answer 19: You can go into GEE and do a search for DEM. There may be one available for your ROI that is better than SRTM. You can also upload your own DEM to GEE by going to the left window and selecting "Assets" and then the red "NEW" button. This will give you the option to upload either a raster (Geotiff) or a vector file.

Question 20: Can we mask the vector file here in Google Earth Engine to limit our analysis within precise ROI boundaries?

Answer 20: Yes, I believe so. This is similar to the time series demo where we created an extra mask and stats within that demo. You can also upload a vector file to GEE.

Question 21: Is there literature you can connect us with about the smoothing filter and how it changes the data. How does the loss of spatial detail affect your study?

Answer 21:

Here is an article that outlines smoothing approaches for land cover classification, however I believe it focuses on MODIS data:

https://pasteur.epa.gov/uploads/353/Time-Series%20Smoothing_rse.2016.pdf



Forest Mapping and Monitoring with SAR Data

May 12 - 21

Here is another: <https://ethz.ch/content/dam/ethz/special-interest/baug/igp/photogrammetry-remote-sensing-dam/documents/pdf/tgrs-2col.pdf>

Here is one on speckle from ESA:

https://earth.esa.int/documents/653194/656796/Speckle_Filtering.pdf

Question 22: What is the best polarization for agriculture monitoring?

Answer 22: This depends on the type (structure) of crop, the growth stage of the crop, and the plantation orientation in relation to the view direction of the radar.

Please see Part Two: Exploiting SAR to Monitor Agriculture from the SAR training

Advanced Webinar: SAR for Landcover Applications,

<https://arset.gsfc.nasa.gov/disasters/webinars/2019-SAR>

Question 23: Is there a minimum size for a training polygon? What if the training samples are known. How can you bring them in?

Answer 23: There is no set rule for minimum training polygon but I would say that you would want a minimum polygon size that should contain at least 10 pixels. You should have on the order of 200-300 pixels to train each class.

You can import training samples into GEE through the “Assets” tab in the upper left window and then selecting the red “NEW” button and selecting either vector or raster to upload your own training polygons.

Question 24: How can you be sure that those dark areas are bare soil, and not, for example, water?

Answer 24: That is a very good question. You cannot be entirely sure based on just the radar and that is where the optical data helps to define your classes. In that case you overlay the optical image to verify if in fact those dark areas are open water or fields with no vegetation.

Question 25: I mean how much training data is sufficient. Are there any rules? Does it depend on the application?

Answer 25: There is no set rule however, having at least 200-300 pixels per class is sufficient. You want to have enough training pixels to have an adequate statistical representation of the backscatter characteristics for each class.

Question 26: Can we estimate above-ground tree biomass using google earth engine SAR data?



Answer 26: I believe you can. The SERVIR Handbook has a chapter (#5) on how to estimate above ground biomass, which you might be able to adapt into GEE:

<https://gis1.servirglobal.net/TrainingMaterials/SAR/ch5.pdf>

Question 27: Pixel values for optical imagery are Reflectance. Which value is displayed in radar images in db?

Answer 27: The radar images are displayed in decibels, which are power units.

Question 28: What are/can be the criteria to decide which polarisation (HH/HV/VH/VV) to use for classification, when our focus is on vegetation? Choosing one over the other will affect the accuracy of classification?

Answer 28: HV is better for detecting vegetation however, both bands contain information that is complementary and both should be used. The good thing about Random Forest is that you do not run into the problem of degrading your results when using multiple input bands since the algorithm does not overfit. However, I suggest running a classification with one polarization and then both and compare the results.

Question 29: Why don't you specify randomforest() parameters like the number of trees, brag fraction, etc. ?

Answer 29: This demonstration was just to get you started. You can play with the parameters to adjust your results. For anyone wanting to learn more about the different parameters that are part of Random Forest, go to the "Docs" tab in the upper left window and type `ee.Classifier.randomForest` and you will see all the options (e.g. *numberOfTrees*, *variablesPerSplit*, *minLeafPopulation*, *bagFraction*, *outOfBagMode*)

Question 30: As you select pixels for a training set, do you want to pick ones that appear to be 'pure?' That is they are visually homogeneous. Or, If you know a location is water, forest, etc., but displays some spectral variation, like water with different surface smoothness, glint, etc. that is, is it better to put all types of water in a single training set or create several different one. For example one for high glint water, one for high roughness water, separate low glint and low roughness, so they are classified separately and combine them into a single "water" classification, a posteriori?

Answer 30: I like to have my classes relatively pure. If you have rough water, you might want to call it Rough water as a training class. Then you can combine those in the final classification.



Question 31: Can the same code be run for high-resolution satellite data such as LISS IV? and does Google Earth Engine allow access to LISS IV?

Answer 31: I am not familiar with this dataset. You can search within GEE and see if that dataset is available.

Question 32: Is there any book available where I can find different spectral reflection /backscattered radiation from different features available on the earth?

Answer 32: The backscatter characteristics of a feature/object on the surface is not as unique as spectral reflectance and hence you will not find a library with the backscatter characteristics of features. For example, backscatter of a forest might vary in different parts of the world depending on many variables such as its structure, roughness, moisture, etc. You might find general backscatter values for a forest but the range might be quite large.

Question 33: How can we do canopy moisture retrieval using SAR data?

Answer 33: There are studies that have done this but it is difficult to estimate canopy moisture with SAR. The band used is important since the L-band tends to penetrate through the vegetation and the moisture estimate will not necessarily be from the canopy surface of the vegetation. C-band does not tend to penetrate as much as L-band and estimation of canopy moisture (with C-band) is probably more related to the canopy of the vegetation. Here are some references:

<http://users.monash.edu.au/~jpwalker/papers/tgars15-4.pdf>

Professor Alexandra Konings of Stanford University has several publications on this topic:

<https://koningslab.stanford.edu/publications>

Question 34: Can you point me to a method to avoid (or assess) autocorrelation issues between training data and validation data?

Answer 34: The following page contains a discussion of autocorrelation issues and a sample code at the bottom of the page:

<https://developers.google.com/earth-engine/classification>

Question 35: Other than treating each data separately is it possible to make a fusion of both sentinel 1 and landsat 8 and then apply the classification? If yes, how? Which polarization to choose? And what is the extra information to expect out of the fusion?



Forest Mapping and Monitoring with SAR Data

May 12 - 21

Answer 35: There are publications out there that have developed algorithms/methodologies with fusion. I suggest you explore those methods as well as the fusion of each polarization to test which one results in higher classification accuracy results or you can also create a HH/HV or VV/VH ratio and use that as the layer to fuse with Landsat data.

Question 36: Can one run an unsupervised classification in GEE?

Answer 36: Yes, you can. Here is a link with more information:

<https://developers.google.com/earth-engine/clustering>

Question 37: How will you export the classified map for further reference and map composition?

Answer 37: The last portion of the demo has the code to export the classification results as a GeoTIFF image.

Question 38: May I be able to import a shapefile with the training/validation data into GEE?

Answer 38: Yes you can. Please refer to question 8.

Question 39: Can we do object based classification in SAR data ?

Answer 39: I believe you can, though I have never tried it in GEE. There is discussion of object based classification in the following discussing forum :

<https://gis.stackexchange.com/questions/273658/performing-object-based-image-classification-in-google-earth-engine>

Question 40: Can SAR data could be used as an alternative for LIDAR, for calculating tree height?

Answer 40: LIDAR measures tree height with higher accuracy than SAR. I would say that SAR can be used as an alternative to LIDAR - since LIDAR is considered the reference dataset. The last session on May 21 will focus on forest height and there will be comparisons with LiDAR.

Question 41: Any references you can suggest about Random Forest?

Answer 41: Here is the original description of Random Forest:

https://www.stat.berkeley.edu/~breiman/RandomForests/cc_home.htm



Forest Mapping and Monitoring with SAR Data

May 12 - 21

And here is a very nice video with an explanation of a number of different classification algorithms in GEE, including Random Forest:

https://www.youtube.com/watch?v=D_KaouS3q20

Question 42: Is it possible to determine forest fire using these methods?

Answer 42: Yes. If you train accordingly, you can apply it.

Question 43: How specific can we be classifying crops versus biomass.

Answer 43: I am not sure what is meant by crops vs. biomass, since these are two very different parameters. We conducted a SAR webinar for Agriculture that discusses how to identify different types of crops with SAR. The SAR Handbook from NASA SERVIR is also a great resource that has examples of how to calculate vegetation biomass with SAR.

Question 44: Does accuracy vary if we use different polygons than training data?

Answer 44: Yes it will vary. A true accuracy assessment should be done with pixels that are not part of the training data.

Question 45: Which band has better penetration through the dense canopy.

Answer 45: The longer the wavelength the greater the penetration through the canopy. P-band would have the greatest penetration, followed by L-band and then C-band. L-band data from PALSAR and C-band data from Sentinel-1 are freely available.

Question 46: Why is Band 1 needed for vegetation classification?

Answer 46: Band 1 of Landsat 8 is not really used for vegetation classification, since it is for aerosols/clouds. However, sometimes it helps to include band 1, to avoid confusion in the band numbers when analyzing imagery.

Question 47: For both Ls8 and SAR for classifiers, can you also down-sample by setting the scale to 10?

Answer 47: Yes you can.

Question 48: What is the effect of the coregistration between these collections? How are you considering this effect?

Answer 48: The coregistration is not perfect and might be 1-2 pixels off in some places. In this example I did not see any coregistration offset between the images.



Forest Mapping and Monitoring with SAR Data

May 12 - 21

If there were coregistration offsets that were significant then you can always perform a correction through manual coregistration using other software such as ENVI.